Mary Ellen Zurko
Associate Editor in Chief

Unusable Security for Attackers

One of the things that makes security research different from other research is the presence of attackers, potentially or in actuality. The early research I was exposed to barely touched on the attacker. The Trusted Computer System Evaluation Criteria from the 1980s had hardly a whisper of functionality specifically for countering attacks, beyond auditing security relevant events. When we were researching and composing secure systems, most of us thought that antivirus, when it emerged, was a fool's game. Who could possibly catch all the different ways an attacker might go about breaching a system? The first question put to a presentation on intrusion detection system (IDS) research was predictably "How did you know that the system was free of attacks when you baselined it?"

The earliest usable security research had no concern for attackers. It emphasized the modest task of just making security mechanisms effectively usable by well-intentioned users. Such research focused on passwords and encryption, in structured lab scenarios, or through qualitative interviews and surveys of organizational behavior. I ended up publishing the first usable security paper that had an attacker model, and also measured user behavior in context, outside of the lab.[1] I was trying to prove to management that our plans for locking down a user feature were insufficient for real use because of the user interactions required should a potential attack attempt to subvert it. Insecure user behavior was not as prevalent in my experiment as I anticipated. (I had a dim view of making users do anything to keep themselves secure). However, the demonstration of such user behavior, in the best possible ecologically valid circumstances, showed that there was a number of users who would click through two dialogs allowing potentially malicious code to run. This garnered support for additional features to lock down the permissions restricting active content in e-mail in the product I was working on.

©SHUTTERSTOCK.COM/VS148

Including attacker considerations when trying to research or develop security solutions is hard for all sorts of reasons. While it turned out I followed ethical principles in my study, it was in a context without an Institutional Review Board (IRB), and I do not recall any discussion around ethics or IRBs in security research at the time. Similarly, in current discussions of what makes a research paper an exemplary candidate for the annual Best Scientific Cybersecurity Research award, none of the specific guidelines on what might make it a strong scientific contribution are about attackers or attacks per se. To my mind, they might be applied to all sorts of research to evaluate its scientific aspect, which leaves a gap in what might be unique to *scientific cybersecurity*. As a cybersecurity professional, I always want to look more closely at gaps, and I see that gap as scientific examination of attackers and their attacks.

After early IDS work, the first research I recall seeing focusing on the attacker themselves was a deception paper at New Security Paradigms Workshop.[2] Bob Blakley had typically evocative thoughts on the topic, such as hiring a magician in your cybersecurity

Digital Object Identifier 10.1109/MSEC.2023.3308648
Date of current version: 10 November 2023

team, and turning around your design thought process to focus an authentication experience that is unusable specifically to attackers. Honey passwords seem now to be the natural outgrowth of that thought experiment. Elie Bursztein of Google might be the first admitted magician in our cybersecurity ranks. And "oppositional human factors" is a phrase that has been introduced into the research literature. We've come a long way since then.

Recent cybersecurity research on attackers tends to be about the beneficial attackers. I'm excited and fascinated every time a new one comes out. Votipka et al.[3] interviewed those two types of attackers. Software security testers focus on finding vulnerabilities before a release, while white-hat hackers focus on doing so after a release. The study provides an example of thoughtful recruitment and screening of those difficult-to-get populations, careful qualitative analysis of their interview data, and a number of findings that provide insights into those attackers (as well as their processes). For example, both experience with vulnerability discovery and knowledge about the underlying system were found to be critical to those attackers' success. Their hacker interviewees specifically attempted to maximize value, highlighting how important the goal of any attacker is to their strategies. Omer et al.[4] combines surveys and interviews to get at bug hunter motivation as well as pertinent aspects of bug-bounty programs and platforms. Focusing on the motivations of this type of attacker, monetary rewards and how they are determined and managed are discussed (echoing the hackers in the 2018 study). The importance of learning, community, and technology familiarity (another validation of the 2018 study) are factors for these attackers. As is the likelihood of finding a bug. Other research, such as Nosco et al.[5] and Alomar et al.[6] has focused on optimizing processes for teams of vulnerability finders and responses to reported vulnerability discoveries, respectively.

In contrast to the work on beneficial hackers, using qualitative studies or training scenarios, I found the Tularosa study focusing on red-team attackers a bit of an eye opener.[7] The authors ran a controlled experiment designed to understand how the use of decoys for deceptive defense in the cyber realm impacted both cyber-based attack tactics, techniques, and procedures (TTPs) and self-reported measurements of psychological state. They hired more than 130 professional red teamers, brought them to a cyber range for a two-day event, and gave them the open-ended task of conducting recon on the network, and locating vulnerable services, misconfigurations, and working exploits. They ran subsets of the red teamers against different conditions across the two days, with and without the presence of decoys, and being disclosed on the potential presence of decoys (or not). They found that the combination of both the actual presence of decoys, and the attackers being told that deception is present, had the greatest impact on their cyberattack behavior. Subsequent work analyzes the red-teamers self-reported emotional state, mapping it to the likelihood of a transition between recon, intrusion, and exploit events, as measured during that two-day event.[8]

There are a handful of other researchers in the literature approaching the challenge of measuring attackers, with their attacks, and drawing inferences about the effectiveness of defenses aimed directly at the human attackers, such as decoys and honey pots. One combines a personality test, a cyber expertise test, and a capture the flag (CTF) event with honeypot to log and categorize the commands used by each attacker, and to find correlations across measured

Digital Object Identifier 10.1109/MSEC.2023.3315882

personality, tested expertise, and those CTF actions.[9] From a scientific process point of view, the experimental choices that need to be made impact both the external validity and the variables that can be measured of such experiments are almost dauntingly legion. Starting with focusing on attacker participants, one of the earliest studies of how attackers think was a qualitative study that developed attacker mental models.[10] It used an author's position in the attacker community to both convenience-sample the participants and to check the quality of their reputation with others in that community. Not a very satisfying approach from the point of view of replicating such an experiment, though an impressive initial baseline likely to have some ecological validity. Hired professional red teamers such as those in the Tularosa study certainly seem closer to the demographics of the most effective attackers in the wild than students learning attack techniques or MTurkers with self-reported experience. However, a range of attackers exist in the wild. Defenses aimed directly at the human attackers may vary in efficacy based on aspects such as attacker experience, goals, or measurable personality traits. Attempting to simplify how attacker participants might be selected, another study measured whether technical professionals with adjacent IT experience might proxy for effective attackers.[11] They were given some targeted training and a very specific goal in a highly structured cyber range. Even in the control condition without deception, these attackers-in-training were only 63% successful, leading the authors to recommend sticking with participants with attack experience for such experiments.

How might subsequent studies select attacker types to focus on? Script kiddies, petty thieves, virus writers, professional criminals, and government agents (to quote one study) all seem wildly different in terms of experience, skills, and goals. They may also be very different in terms of personality and other behavioral and psychological reactions. Research that measures the attacker directly almost invariably includes the Big Five Inventory that measures extraversion, agreeableness, conscientiousness, neuroticism, and openness. Research aimed at insider attackers measures the Dark Triad of narcissism, psychopathy, and Machiavellianism. Studies measure stress through a survey or physiological data. Surveys asking about confusion, frustration, self-doubt, confidence, and surprise are common when deception is part of the task. Homogeneous participant populations give a point of reference for all these measurements, but it does not yet seem clear how to map them to the categories often used for attackers in the wild.

There is a myriad of ecological concerns in an experiment's attack task itself, which will be familiar to red teamers and trainers, but are largely new to the usable security research community that focuses on measuring humans and their interactions. For external validity, the attack task should match the attacker type, and use tools the attacker is familiar with or likely to try out. Different parts of the cyber kill chain, different TTPs, will match with different tools and experimental tasks. "Living off the land" is a popular experimental configuration since it minimizes attack-specific tools. This places additional focus on ensuring the experimental cyber testbed is realistically configured as a target of interest, which leads to a large number of details that need to be thoughtfully configured and documented, including system type and scale; account types, credentials, and access; network configuration; and decisions made about the security mechanisms and hygiene that the attacker must navigate.

Are there alternatives to this complex configuration? Can we learn something about attackers from attack tasks structured more simply, as in training? What about the parts of the kill chain where the attacker is not interacting with a specific potential target? What if the attacker is asked to provide a typical goal themselves, or given an open-ended situation and asked how they would proceed? There's clearly a role to play for more focused studies to determine which of the many experimental conditions matter. Both positive and negative results can contribute in this area, illuminating the impacts of tradeoffs that allow for more experimental control but lack external validity. Such experiments might not always pass the "shiny innovation" sniff test that sometimes pervades program committees. But they are essential to progress in more rigorous and scientific understanding of defending against the attackers behind the attacks. ■

Acknowledgment

This work was supported by the Intelligence Advanced Research Projects Activity (IARPA), under the ReSCIND program.

Distribution Statement A. Approved for public release. Distribution is unlimited. This material is based upon work supported under Air Force Contract FA8702-15-D-0001. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of United States Air Force or the Intelligence Advanced Research Projects Activity (IARPA).

References

1. M. E. Zurko, C. Kaufman, K. Spanbauer, and C. Bassett, "Did you ever have to make up your mind? What notes users do when faced with a security decision," in *Proc. 18th Annu. Comput. Secur. Appl.*

Conf., Las Vegas, NV, USA, 2002, pp. 371–381, doi: 10.1109/CSAC.2002.1176309.
2. M. H. Almeshekah and E. H. Spafford, "Planning and integrating deception into computer security defenses," in *Proc. New Secur. Paradigms Workshop*, 2014, pp. 127–138, doi: 10.1145/2683467.2683482.
3. D. Votipka, R. Stevens, E. Redmiles, J. Hu, and M. Mazurek, "Hackers vs. testers: A comparison of software vulnerability discovery processes," in *Proc. IEEE Symp. Secur. Privacy*, San Francisco, CA, USA, 2018, pp. 374–391, doi: 10.1109/SP.2018.00003.
4. O. Akgul et al., "Bug hunters' perspectives on the challenges and benefits of the bug bounty ecosystem," in *Proc. 32nd USENIX Secur. Symp.*, 2023, vol. 2301, pp. 1–23.
5. T. Nosco et al., "The industrial age of hacking," in *Proc. 29th USENIX Secur. Symp.*, 2020, pp. 1129–1146. [Online]. Available: https://www.usenix.org/conference/usenixsecurity20/presentation/nosco
6. N. Alomar, P. Wijesekera, E. Qiu, and S. Egelman, "'You've got your nice list of bugs, now what?' Vulnerability discovery and management processes in the wild," in *Proc. 16th Symp. Usable Privacy Secur. (SOUPS)*, 2020, pp. 319–339. [Online]. Available: https://www.usenix.org/conference/soups2020/presentation/alomar
7. K. Ferguson-Walter, M. Major, C. K. Johnson, and D. H. Muhleman, "Examining the efficacy of decoy-based and psychological cyber deception," in *Proc. 30th USENIX Secur. Symp.*, 2021, pp. 1127–1144.
8. R. Gabrys et al., "Emotional state classification and related behaviors among cyber attackers," in *Proc. 56th Hawaii Int. Conf. Syst. Sci.*, 2023. [Online]. Available: https://hdl.handle.net/10125/102735
9. M. Odemis, C. Yucel, and A. Koltuksuz, "Detecting user behavior in cyber threat intelligence: Development of Honeypsy system," *Secur. Commun. Netw.*, vol. 2022, Jan. 2022, Art. no. 7620125, doi: 10.1155/2022/7620125.
10. T. C. Summers, K. J. Lyytinen, T. Lingham, and E. A. Pierce, "How hackers think: A study of cybersecurity experts and their mental models," in *SSRN Electron. J.*, Jan. 2013. [Online]. Available: https://ssrn.com/abstract=2326634
11. T. Shade, A. Rogers, K. J. Ferguson-Walter, S. B. Elsen, D. Fayette, and K. E. Heckman, "The moonraker study: An experimental evaluation of host-based deception," in *Proc. 53rd Hawaii Int. Conf. Syst. Sci.*, 2020, pp. 1875–1884.

Digital Object Identifier 10.1109/MSEC.2023.3324567